```
from google.colab import files
f=files.upload()
```

```
Choose Files dataset_Facebook.csv
• dataset_Facebook.csv(text/csv) - 37891 bytes, last modified: 5/18/2023 - 100% done
Saving dataset_Facebook.csv to dataset_Facebook.csv
```

```
import pandas as pd
import numpy as np
```

```
df=pd.read_csv('dataset_Facebook.csv',sep=';')
df
```

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lif Post Impres
0	139441	Photo	2	12	4	3	0.0	2752	
1	139441	Status	2	12	3	10	0.0	10460	
2	139441	Photo	3	12	3	3	0.0	2413	
3	139441	Photo	2	12	2	10	1.0	50128	
4	139441	Photo	2	12	2	3	0.0	7244	
...	...	...	...	...	...	...	...	...	
495	85093	Photo	3	1	7	2	0.0	4684	
496	81370	Photo	2	1	5	8	0.0	3480	
497	81370	Photo	1	1	5	2	0.0	3778	
498	81370	Photo	3	1	4	11	0.0	4156	
499	81370	Photo	2	1	4	4	NaN	4188	

500 rows × 19 columns

```
df.head()
```

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifet Post To Impressi
0	139441	Photo	2	12	4	3	0.0	2752	5
1	139441	Status	2	12	3	10	0.0	10460	19
2	139441	Photo	3	12	3	3	0.0	2413	4
3	139441	Photo	2	12	2	10	1.0	50128	87
4	139441	Photo	2	12	2	3	0.0	7244	13

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 19 columns):
 #   Column                                                    Non-Null Count  Dtype
---  ------                                                    --------------  -----
 0   Page total likes                                          500 non-null    int64
 1   Type                                                      500 non-null    object
 2   Category                                                  500 non-null    int64
 3   Post Month                                                500 non-null    int64
 4   Post Weekday                                              500 non-null    int64
 5   Post Hour                                                 500 non-null    int64
 6   Paid                                                      499 non-null    float64
 7   Lifetime Post Total Reach                                 500 non-null    int64
 8   Lifetime Post Total Impressions                           500 non-null    int64
 9   Lifetime Engaged Users                                    500 non-null    int64
 10  Lifetime Post Consumers                                   500 non-null    int64
 11  Lifetime Post Consumptions                                500 non-null    int64
 12  Lifetime Post Impressions by people who have liked your Page  500 non-null  int64
 13  Lifetime Post reach by people who like your Page          500 non-null    int64
 14  Lifetime People who have liked your Page and engaged with your post  500 non-null  int64
 15  comment                                                   500 non-null    int64
 16  like                                                      499 non-null    float64
 17  share                                                     496 non-null    float64
 18  Total Interactions                                        500 non-null    int64
dtypes: float64(3), int64(15), object(1)
memory usage: 74.3+ KB
```

```
df.isnull()
```

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users
0	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...
495	False	False	False	False	False	False	False	False	False	False
496	False	False	False	False	False	False	False	False	False	False
497	False	False	False	False	False	False	False	False	False	False
498	False	False	False	False	False	False	False	False	False	False
499	False	False	False	False	False	False	True	False	False	False

500 rows × 19 columns

```
df.isnull().sum()
```

```
Page total likes;Type;Category;Post Month;Post Weekday;Post Hour;Paid;Lifetime Post Total Reach;Lifetime Post Total
Impressions;Lifetime Engaged Users;Lifetime Post Consumers;Lifetime Post Consumptions;Lifetime Post Impressions by people who have
liked your Page;Lifetime Post reach by people who like your Page;Lifetime People who have liked your Page and engaged with your
post;comment;like;share;Total Interactions    0
dtype: int64
```

```
df.nunique()
```

```
Page total likes;Type;Category;Post Month;Post Weekday;Post Hour;Paid;Lifetime Post Total Reach;Lifetime Post Total
Impressions;Lifetime Engaged Users;Lifetime Post Consumers;Lifetime Post Consumptions;Lifetime Post Impressions by people who have
liked your Page;Lifetime Post reach by people who like your Page;Lifetime People who have liked your Page and engaged with your
post;comment;like;share;Total Interactions    500
dtype: int64
```

```
df.dropna(how='any' , axis=0)
df
```

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lif Post Impres
0	139441	Photo	2	12	4	3	0.0	2752	
1	139441	Status	2	12	3	10	0.0	10460	
2	139441	Photo	3	12	3	3	0.0	2413	
3	139441	Photo	2	12	2	10	1.0	50128	
4	139441	Photo	2	12	2	3	0.0	7244	
...	...	...	...	...	...	...	...	...	
495	85093	Photo	3	1	7	2	0.0	4684	
496	81370	Photo	2	1	5	8	0.0	3480	
497	81370	Photo	1	1	5	2	0.0	3778	
498	81370	Photo	3	1	4	11	0.0	4156	
499	81370	Photo	2	1	4	4	NaN	4188	

500 rows × 19 columns

1)Creating Data Subsets

```
df1=df[['Category','Type','Lifetime Post Total Reach','Total Interactions']].loc[1:245]
df1
```

	Category	Type	Lifetime Post Total Reach	Total Interactions
1	2	Status	10460	164
2	3	Photo	2413	80
3	2	Photo	50128	1777
4	2	Photo	7244	393
5	2	Status	10472	186
...	...	...	...	...
241	1	Photo	4892	144
242	2	Status	17360	572
243	1	Video	21872	409
244	2	Photo	180480	6334
245	1	Photo	44464	188

```
df2=df[['Post Month','Post Weekday','Post Hour','Lifetime Post Consumers']].loc[245:500]
df2
```

	Post Month	Post Weekday	Post Hour	Lifetime Post Consumers
245	7	2	13	930
246	7	2	8	503
247	7	1	12	482
248	7	1	6	512
249	7	7	11	444
...	...	...	...	...
495	1	7	2	708
496	1	5	8	508
498	1	4	11	574

2)Merging Dataset

```
df_merge=pd.concat([df1,df2])
df_merge
```

	Category	Type	Lifetime Post Total Reach	Total Interactions	Post Month	Post Weekday	Post Hour	Lifet P Consum
1	2.0	Status	10460.0	164.0	NaN	NaN	NaN	N
2	3.0	Photo	2413.0	80.0	NaN	NaN	NaN	N
3	2.0	Photo	50128.0	1777.0	NaN	NaN	NaN	N
4	2.0	Photo	7244.0	393.0	NaN	NaN	NaN	N
5	2.0	Status	10472.0	186.0	NaN	NaN	NaN	N
...	...	...	...	...	...	...	...	
495	NaN	NaN	NaN	NaN	1.0	7.0	2.0	7(
496	NaN	NaN	NaN	NaN	1.0	5.0	8.0	5(
497	NaN	NaN	NaN	NaN	1.0	5.0	2.0	5

3)Transposing Data

```
df.transpose()
```

	0	1	2	3	4	5	6	
Page total likes	139441	139441	139441	139441	139441	139441	139441	139
Type	Photo	Status	Photo	Photo	Photo	Status	Photo	Pl
Category	2	2	3	2	2	2	3	
Post Month	12	12	12	12	12	12	12	
Post Weekday	4	3	3	2	2	1	1	
Post Hour	3	10	3	10	3	9	3	
Paid	0.0	0.0	0.0	1.0	0.0	0.0	1.0	
Lifetime Post								

```
df1.transpose()
```

	1	2	3	4	5	6	7	8	9
Category	2	3	2	2	2	3	3	2	3
Type	Status	Photo	Photo	Photo	Status	Photo	Photo	Status	Photo
Lifetime Post Total Reach	10460	2413	50128	7244	10472	11692	13720	11844	4694

Consumptions

```
df2.transpose()
```

	245	246	247	248	249	250	251	252	253	254	...	490	49
Post Month	7	7	7	7	7	7	7	7	7	7	...	1	
Post Weekday	2	2	1	1	7	7	6	6	5	5	...	6	
Post Hour	13	8	12	6	11	6	11	5	12	3	...	11	
Lifetime													

4)Sorting Data

```
df.sort_values(by='Category')
df.sort_index()
```

	Page total	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total	Lif Post

5)Shape and Reshape of Data

```
df.shape
```

```
(500, 19)
```

1	139441	Status	2	12	3	10	0.0	10460	

```
pivot_table=pd.pivot_table(df,index=['Type', 'Category'], values='comment')
pivot_table
```

		comment
Type	**Category**	
Link	**1**	2.900000
	2	2.000000
	3	2.000000
Photo	**1**	5.897297
	2	11.692308
	3	6.913333
Status	**1**	4.333333
	2	9.921053
	3	2.750000

✓ 0s completed at 11:12 AM